Quantitative study of single molecule location estimation techniques

Anish V. Abraham[1,2]**, Sripad Ram**[1]**, Jerry Chao**[1,2]**, E. S. Ward**[1]**, and Raimund J. Ober**[1,2]

[1]*Department of Immunology, University of Texas Southwestern Medical Center, 6000 Harry Hines Boulevard MC9093 Dallas, TX 75390, U.S.A.*
[2]*Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75080, U.S.A.*

ober@utdallas.edu

Abstract: Estimating the location of single molecules from microscopy images is a key step in many quantitative single molecule data analysis techniques. Different algorithms have been advocated for the fitting of single molecule data, particularly the nonlinear least squares and maximum likelihood estimators. Comparisons were carried out to assess the performance of these two algorithms in different scenarios. Our results show that both estimators, on average, are able to recover the true location of the single molecule in all scenarios we examined. However, in the absence of modeling inaccuracies and low noise levels, the maximum likelihood estimator is more accurate than the nonlinear least squares estimator, as measured by the standard deviations of its estimates, and attains the best possible accuracy achievable for the sets of imaging and experimental conditions that were tested. Although neither algorithm is consistently superior to the other in the presence of modeling inaccuracies or misspecifications, the maximum likelihood algorithm emerges as a robust estimator producing results with consistent accuracy across various model mismatches and misspecifications. At high noise levels, relative to the signal from the point source, neither algorithm has a clear accuracy advantage over the other. Comparisons were also carried out for two localization accuracy measures derived previously. Software packages with user-friendly graphical interfaces developed for single molecule location estimation (EstimationTool) and limit of the localization accuracy calculations (FandPLimitTool) are also discussed.

© 2009 Optical Society of America

OCIS codes: (100.2960) Image analysis; (180.2520) Fluorescence microscopy.

References and links

1. X. S. Xie, P. J. Choi, G. W. Li, N. K. Lee, and G. Lia, "Single-molecule approach to molecular biology in living bacterial cells," Annu. Rev. Biophys. **37**, 417–444 (2008).
2. W. E. Moerner, "New directions in single-molecule imaging and analysis," Proc. Natl. Acad. Sci. U.S.A. **104**, 12596–12602 (2007).
3. M. Dahan, S. Levi, C. Luccardini, P. Rostaing, B. Riveau, and A. Triller, "Diffusion dynamics of glycine receptors revealed by single-quantum dot tracking," Science. **302**, 442–445 (2003).
4. P. H. M. Lommerse, G. A. Blab, L. Cognet, G. S. Harms, B. E. Snaar-Jagalska, H. P. Spaink, and T. Schmidt, "Single-molecule imaging of the H-Ras membrane-anchor reveals domains in the cytoplasmic leaflet of the cell membrane," Biophys. J. **86**, 609–616 (2004).

5. P. H. M. Lommerse, B. E. Snaar-Jagalska, H. P. Spaink, and T. Schmidt, "Single-molecule diffusion measurements of H-Ras at the plasma membrane of live cells reveal microdomain localization upon activation," J. Cell. Sci. **118**, 1799–1809 (2005).

6. K. Murase, T. Fujiwara, Y. Umemura, K. Suzuki, R. Iino, H. Yamashita, M. Saito, H. Murakoshi, K. Ritchie, and A. Kusumi, "Ultrafine membrane compartments for molecular diffusion as revealed by single molecule techniques," Biophys. J. **86**, 4075–4093 (2004).

7. M. J. Rust, M. Bates, X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," Nat. Meth. **3**, 793–796 (2006).

8. B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," Science. **319**, 810–813 (2008).

9. S. T. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," Biophys. J. **91**, 4258–4272 (2006).

10. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," Science. **313**, 1642–1645 (2006).

11. M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," Biophys. J. **81**, 2378–2388 (2001).

12. S. M. Kay, *Fundamentals of statistical signal processing* (Prentice Hall, 1993).

13. R. J. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single-molecule microscopy," Biophys. J. **86**, 1185–1200 (2004).

14. S. Ram, E. S. Ward, and R. J. Ober, "A stochastic analysis of performance limits for optical microscopes," Multidim. Syst. Sign. Process. **17**, 27–57 (2006).

15. J. Markham and J. A. Conchello, "Fast maximum-likelihood image-restoration algorithms for three-dimensional fluorescence microscopy," J. Opt. Soc. Am. **18**, 1062–1071 (2001).

16. P. J. Verveer and T. M. Jovin, "Efficient superresolution restoration algorithms using maximum a posteriori estimations with application to fluorescence microscopy," J. Opt. Soc. Am. **14**, 1696–1706 (1997).

17. R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," Biophys. J. **82**, 2775–2783 (2002).

18. A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin, "Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization," Science. **300**, 2061–2065 (2003).

19. X. Qu, D. Wu, L. Mets, and N. F. Scherer, "Nanometer-localized multiple single-molecule fluorescence microscopy," Proc. Natl. Acad. Sci. U.S.A. **101**, 11298–11303 (2004).

20. H. Park, G. T. Hanson, S. R. Duff, and P. R. Selvin, "Nanometer localization of single ReAsH molecules," J. Microsc. **216**, 199–205 (2004).

21. J. H. Kim and R. G. Larson, "Single-molecule analysis of 1D diffusion and transcription elongation of T7 RNA polymerase along individual stretched DNA molecules," Nucleic Acids Res. **35**, 3848–3858 (2007).

22. S. Ram, E. S. Ward, and R. J. Ober, "How accurately can a single molecule be localized in three dimensions using a fluorescence microscope?" Proc. SPIE. **5699**, 426–435 (2005).

23. "EstimationTool," http://www4.utsouthwestern.edu/wardlab/EstimationTool.

24. "FandPLimitTool," http://www4.utsouthwestern.edu/wardlab/FandPLimitTool.

25. M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, UK, 1999).

26. P. Torok and F.-J. Kao, *Optical Imaging and Microscopy* (Springer, 2003).

27. B. Zhang, J. Zerubia, and J. C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," Appl. Opt. **46**, 1819–1829 (2007).

28. J. S. Biteen, M. A. Thompson, N. K. Tselentis, G. R. Bowman, L. Shapiro, and W. E. Moerner, "Super-resolution imaging in live Caulobacter crescentus cells using photoswitchable EYFP," Nat. Methods. **5**, 947–949 (2008).

29. S. Zacks, *Theory of Statistical Inference* (John Wiley and Sons, 1971).

1. Introduction

Within a relatively short period of time, through various advances in experimental and imaging techniques, single molecule experiments have become almost routine and have been used in different areas of biology to elicit important new insights [1, 2]. Such experiments provide access to biological information that would otherwise have been obscured by the averaging effects of bulk fluorescence imaging.

Quantitative analysis plays a crucial role in gaining information from single molecule data, and the first step in many single molecule data analysis techniques is the determination of the location of a single molecule with subpixel accuracy from its image. For example, to study the diffusion of specific proteins, fluorescently tagged proteins are first localized from a series of

consecutive time-lapse images [3–6].

Among the recent advances in imaging technologies has been the development of super-resolution imaging techniques that are based on localization approaches, e.g. [7–10]. These techniques rely on exciting small sparsely distributed subsets of fluorophores. The location of each of these excited fluorophores is then estimated and the super-resolution image is assembled from the estimated locations of the fluorophores. The quality of the reconstructed super-resolution image is dependent on the accuracy with which the location of individual fluoropheres can be determined.

Determining the location of a single molecule from its image is not straightforward. The emission and detection of photons from a single molecule are stochastic processes. Moreover, as there is no ideal noise-free detector, the data is further corrupted by pixelation and readout-noise. As a result of these factors, the location of a single molecule cannot be precisely determined from its image, but only estimated with a certain error.

There are various techniques by which the location of a single molecule can be estimated from its image. The question arises as to which of these techniques produces the most accurate results. Cheezum et. al. [11] have previously carried out a comparative study of some of the different approaches. The authors estimated locations of single molecules by calculating, for example, the center-of-mass, and by fitting Gaussian profiles using the least squares algorithm [11]. They concluded that for point sources, amongst the investigated algorithms, fitting Gaussian profiles using the least squares algorithm produced the most accurate results.

The maximum likelihood estimator is a classical estimator that has been extensively studied and used in statistical and signal processing applications due to its well established performance [12]. For a given data set and underlying probability model, the maximum likelihood estimator returns values for the model parameters which are most likely to produce the observed data. On the other hand, the least squares algorithm works by finding the set of model parameters that produces the least difference, in the least squares sense, between the model and the observed data. Thus the two algorithms use different approaches to solve parametric estimation problems.

Since the maximum likelihood estimator was not investigated in [11] the question arises as to how the maximum likelihood estimator compares with the nonlinear least squares estimator in the context of single molecule localization. It has previously been shown that the maximum likelihood estimator can achieve the best possible accuracy for a few specific scenarios [13, 14]. Reports in the deconvolution literature also show that maximum likelihood based algorithms can produce more accurate results than least squares based algorithms especially for quantum limited data, i.e. Poisson distributed data with low signal levels [15, 16]. Here we further investigate the performance of the maximum likelihood estimator for various estimation scenarios.

There are many techniques other than the ones already mentioned by which the location of a single molecule can be determined. The standard deviation of the location estimates of a single molecule from multiple images is typically taken as a measure of the accuracy of the estimator [13, 17]. The smaller the standard deviation, the better the accuracy of the estimator. The question arises as to what is the best accuracy that can be achieved for a given data set irrespective of the estimation technique used. Having a method to determine this, given only the imaging and experimental conditions, provides two key advantages. One, it allows us to assess the performance of an estimation algorithm, i.e., it allows us to assess how close the accuracy of estimates is to the best possible accuracy. Two, it allows us to make more informed decisions in the design and execution of single molecule imaging experiments with respect to the level of localization accuracy that can be achieved.

Thompson et. al. [17] have previously provided a simple formula which, based on certain approximations, can be used to estimate the localization error for a given set of imaging and experimental conditions. This approach has been extensively used in various studies, see

e.g. [9, 10, 18–21]. On the other hand, discrepancies have been reported between the standard deviation of the nonlinear least squares estimates and the accuracy predicted by this approach, e.g. [7]. There are many factors that could lead to such discrepancies, e.g. inadequate drift correction, optical aberrations. Here, we wanted to examine the validity of the approach itself under various experimental and imaging conditions and explore whether there are inadequacies in the approach that could also contribute to the differences between the experimentally observed and theoretically predicted standard deviations. Our group had previously introduced a framework based on the classical theory of the Fisher information matrix and the Cramer-Rao lower bound to determine the localization accuracy taking into account the stochastic nature of the photon emission and detection processes [13, 14, 22]. The limit of the localization accuracy, or the localization accuracy measure, was defined as the best possible localization accuracy achievable for a given set of imaging and experimental conditions, independent of a specific estimation algorithm [13]. We have shown through simulations that the maximum likelihood estimator can attain the limit of the localization accuracy [13, 14]. Here we compare the limit of the localization accuracy with the estimate of the localization error predicted using the formula provided by Thompson et. al. [17]. We also compare the standard deviations of estimates from both the nonlinear least squares and maximum likelihood estimators with the limit of the localization accuracy.

An important characteristic of an algorithm is its 'robustness', i.e., whether or not it still performs well in the presence of model mismatches and misspecifications. Various examples of model mismatches frequently occur in the analysis of single molecule microscopy data. For computational efficiency Gaussian profiles are often used to fit single molecule images that might be more accurately modeled by Airy profiles. Thus the model used to fit the image may not match the model that best describes the image formation. In fact the analysis in [11] is such that the data is simulated using an Airy profile but is fit with a Gaussian profile. A further example occurs when a parameter, such as the width parameter in the Gaussian or Airy profile, is independently determined. The question then arises how precisely this parameter needs to be known so that the localization algorithm does not suffer in performance. Here we investigate both the maximum likelihood and the nonlinear least squares algorithms with respect to their robustness in the presence of various modeling inaccuracies.

Some of the algorithms to estimate single molecule locations are relatively complex. Similarly, a number of the approaches to calculate the limit of the localization accuracy involve non-trivial computational steps. The lack of appropriate software has sometimes prevented the use of some of these methods. We therefore discuss the EstimationTool [23] and the FandPLimitTool [24], software packages that were developed to facilitate these computations.

2. Materials & Methods

2.1. Simulating single molecule images

To compare the accuracies of the nonlinear least squares and maximum likelihood estimators, we simulated images of a stationary single molecule as acquired by a camera and estimated the location of the single molecule using both estimators. The camera is modeled as a set of pixels. The signal acquired at any pixel is modeled as the sum of the signal from the point source, the signal from the background component, both modeled as independent Poisson random variables, and camera noise modeled as an independent Gaussian random variable [13]. The equations to model the signal at any camera pixel and the equations describing the random variables denoting the various components of the signal are given in the *Appendix A*. The mean of the signal from the point source can be described by either an Airy or a Gaussian photon distribution profile. The equations describing both photon distributions profiles are given in the *Appendix A*. The details of how the single molecule images are generated by realizations

of the Poisson and Gaussian random variables is also described in the *Appendix A*. All calculations involved in generating the single molecule images were performed in MATLAB (The MathWorks, Natick, MA).

2.2. *Fitting single molecule images*

Both the maximum likelihood and nonlinear least squares estimators were used to fit the single molecule images and estimate the location of the single molecule. Both algorithms make use of either the Airy or Gaussian photon distributions profiles mentioned above. The equations describing how the nonlinear least squares and maximum likelihood estimates are calculated are given in *Appendix B*. All calculations involved in estimating the location of a single molecule were performed in MATLAB. For all analyses, only those estimates were used for which the MATLAB optimization functions successfully completed and the estimated location coordinates were within the image (see [23] and *Software* below).

2.3. *Limit of the localization accuracy*

The limit of the localization accuracy provides a lower bound on the standard deviation of estimates from an unbiased estimator for a given set of imaging and experimental conditions [13,14]. For ease of reference we introduce here two notations for the concepts first presented in [13]. The fundamental localization accuracy measure (FLAM) denotes the square root of the Cramer-Rao lower bound on the variance of the 2D location coordinate estimates of the single molecule for the situation when perfect detection conditions are assumed, i.e., a detector of infinite area without pixelation and no extraneous noise sources. In contrast, the practical localization accuracy measure

(PLAM, not to be confused with PALM the high-resolution microscopy technique using photoactivation) refers to the square root of the Cramer-Rao lower bound on the variance of the 2D location coordinate estimates of the single molecule for the practical situation, i.e., when the detection is carried out with a pixelated detector of finite area, and the signal is potentially corrupted by extraneous noise sources such as scattered photons, modeled by an additive Poisson process, and camera readout noise, modeled by an additive Gaussian process.

All calculations for the computation of the limit of the localization accuracy for both the Airy and Gaussian profiles were carried out in MATLAB using custom routines based on the equations in [13,14] (see [24] and *Software* below).

2.4. *Software*

Software packages (EstimationTool and FandPLimitTool) were developed in MATLAB for the estimation of the location of single molecules (EstimationTool) and for the calculations of the limit of the localization accuracy (FandPLimitTool) for a variety of different point spread function models. In the EstimationTool, both the nonlinear least squares and the maximum likelihood algorithms were implemented. Both packages were developed on object-oriented programming methodologies. User-friendly graphical user interfaces have been designed for both packages and are available at `http://www4.utsouthwestern.edu/wardlab`.

3. Results

3.1. *Maximum likelihood estimates have the smallest standard deviation in the ideal case*

Our approach to the assessment of the performance of the nonlinear least squares and maximum likelihood estimators for location estimation was to investigate whether either estimator exhibits bias and to compare the standard deviations of estimates obtained from both estimators. We first considered the ideal case of fitting an accurately specified image profile.
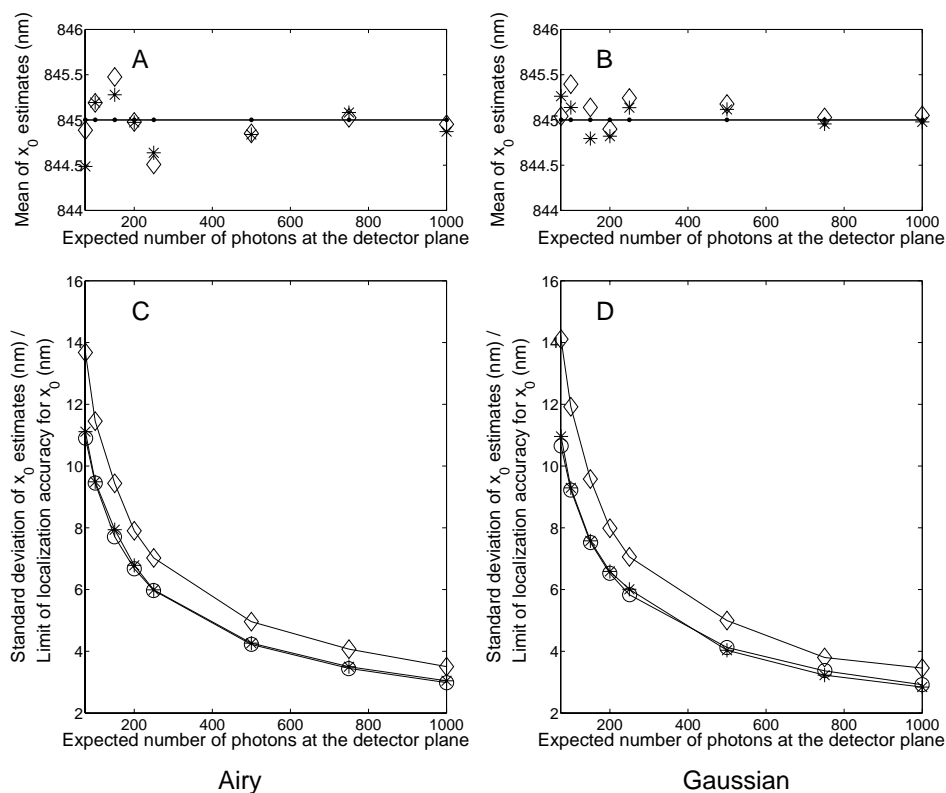
Fig. 1. Comparison of standard deviation of estimates from the nonlinear least squares and maximum likelihood estimators in the ideal case. Panels $A(B)$ and $C(D)$ show the results when the Airy (Gaussian) pixelated profile was used both to generate and fit the single molecule images. Panels $A(C)$ and $B(D)$ show the mean (standard deviations) of the x_0 coordinate estimates from the nonlinear least squares ($\diamond$) and maximum likelihood ($*$) estimators. ($\bullet$) indicates the true x_0 coordinate value and ($\circ$) the PLAM or limit of the localization accuracy of x_0. For each value of the expected number of photons at the detector plane, 1000 images of a stationary single molecule were generated. Location coordinates (x_0, y_0) were estimated from each image using both estimators. Values for the width, photon detection rate, and background parameters, were fixed to the values used to generate the images. The following numerical values were used when simulating the single molecule images. Pixel size: $13\ \mu m \times 13\ \mu m$, pixel array size: 13×13, magnification $M = 100$. The single molecule image was centered within the pixel array. The mean of the background noise component was set to zero. When the Airy pixelated profile was used to generate the single molecule images or to fit them, wavelength $\lambda = 520\ nm$, and numerical aperture $n_a = 1.3$. When the Gaussian pixelated profile was used to generate the single molecule images or to fit them, the width parameter of the Gaussian pixelated profile was calculated as $\sigma = 1.323\lambda/2\pi n_a = 84.225\ nm$.

We considered the two scenarios that are often assumed, one in which the images were simulated using pixelated Gaussian profiles, the other in which they were simulated using pixelated Airy profiles. For each of these two scenarios, sets of one thousand images of a stationary single molecule were simulated for various values of the expected number of photons at the

detector plane. The single molecule location was estimated from each image using both estimators. Here we assumed the absence of any model misspecifications. The type of profile used to fit each image matched the type of profile used to generate that image, i.e., images generated using pixelated Airy profiles were fit using pixelated Airy profiles, and images generated using pixelated Gaussian profiles were fit using pixelated Gaussian profiles. The location coordinates ($\theta := x_0, y_0$) were estimated from each image while the values for the width and photon detection rate parameters were set to the values used to generate the single molecule images.

Examining the location estimates, we first observe that both estimators are, on average, able to accurately recover the true location of the single molecule for each value of the expected number of photons at the detector plane. This was true for both scenarios, i.e., irrespective of whether Airy or Gaussian pixelated profiles were used to generate and fit the data [Fig. 1(a) and 1(b)]. The distribution of the mean of the estimates for each value of the expected photon count with respect to the true location coordinate shows that neither estimator exhibits bias in this scenario.

For each estimator and for each simulation condition, we then calculated the standard deviations of the location estimates. As expected, with increasing photon count the estimates become more accurate. Comparing the standard deviations for the two estimators, we observe that the standard deviations of the estimates from the maximum likelihood estimator are consistently smaller than those from the nonlinear least squares estimator [Fig. 1(c) and 1(d)]. This difference in the standard deviations is consistent across all the expected photon counts that were examined. This indicates that the maximum likelihood estimator is more accurate than the nonlinear least squares estimator in this situation.

3.2. The standard deviations of the maximum likelihood estimates attain the limit of the localization accuracy

It is clear from the previous results (Fig. 1) that the accuracy of single molecule location estimates depends on the choice of the estimation algorithm. The question arises, what is the best possible localization accuracy that can be achieved and how close can we get to that ideal? The limit of the localization accuracy, or PLAM (see *Materials and Methods*), based on the theory of the Fisher information matrix and the Cramer-Rao Lower Bound, provides a lower bound on the variance of estimates obtained using any unbiased estimator [13]. In other words, irrespective of the parameter estimation algorithm used, the limit of the localization accuracy provides a measure of the best possible accuracy with which the location of a single molecule can be estimated for a given set of experimental and imaging conditions. Therefore, we used the limit of the localization accuracy as a standard against which to compare the accuracy of both the nonlinear least squares and maximum likelihood estimators. Calculating the PLAM for each value of the expected number of photons at the detector plane for the imaging and experimental parameters used in Fig. 1, we find that the standard deviations of the estimates from the maximum likelihood estimator match the PLAM [Fig. 1(c) and 1(d)]. This shows that the maximum likelihood estimator achieves the best possible localization accuracy in this scenario, whereas the nonlinear least squares algorithm is suboptimal. These calculations were also repeated by floating the photon detection rate parameter along with the location coordinate parameters and the results were confirmed (data not shown).

3.3. The effect of noise on the standard deviations of estimates

The simulations and estimations performed in the previous comparison were in the absence of extraneous noise (background autofluorescence or readout noise), which is of course not the case in practice. To investigate how both algorithms perform on data with noise, first, we generated sets of one thousand single molecule images for varying levels of additive Poisson
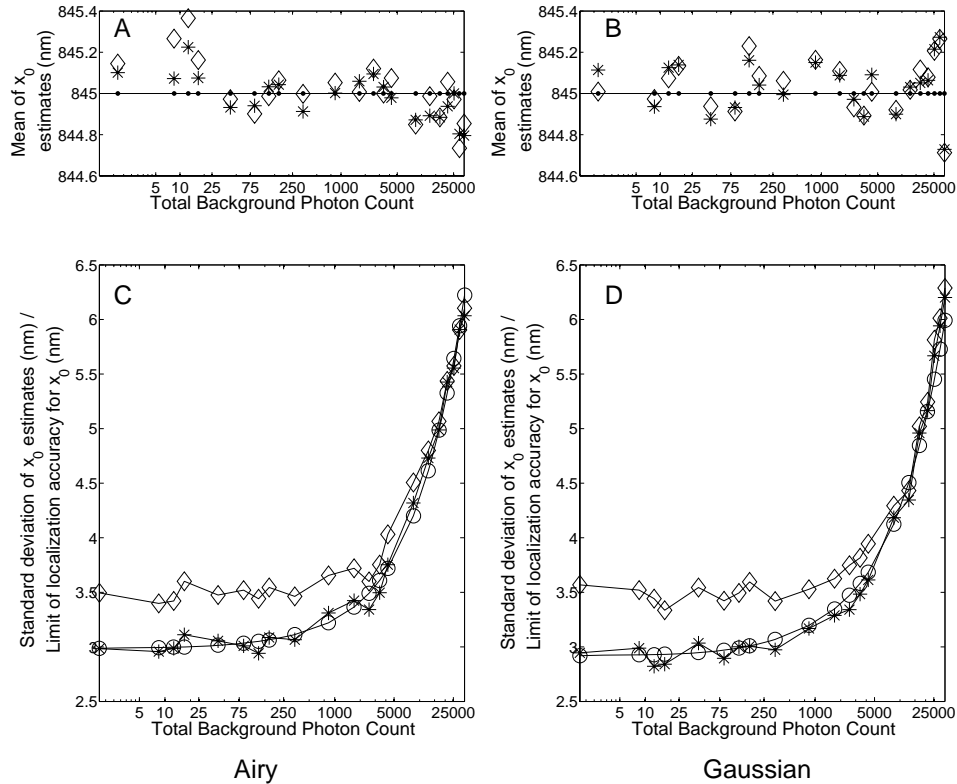
Fig. 2. Comparison of standard deviation of estimates from the nonlinear least squares and maximum likelihood estimators in the presence of additive Poisson noise and absence of model mismatches or misspecifications. Panels A(B) and C(D) show the results when the Airy(Gaussian) pixelated profile was used both to generate and fit the single molecule images. Panels $A(C)$ and $B(D)$ show the mean (standard deviations) of the x_0 coordinate estimates from the nonlinear least squares ($\diamond$) and maximum likelihood ($*$) estimators. ($\bullet$) indicates the true x_0 coordinate value and ($\circ$) the PLAM or limit of the localization accuracy of x_0. For each value of the total background photon count, 1000 images of a stationary single molecule were generated. Location coordinates (x_0, y_0) were estimated from each image using both estimators. Values for the width, photon detection rate, and background parameters, were fixed to the values used to generate the images. The following numerical values were used when simulating the single molecule images. Pixel size: $13 \ \mu m \times 13 \ \mu m$, pixel array size: 13×13, magnification $M = 100$. The single molecule image was centered within the pixel array. When the Airy pixelated profile was used to generate the single molecule images or to fit them, wavelength $\lambda = 520 \ nm$, and numerical aperture $n_a = 1.3$. When the Gaussian pixelated profile was used to generate the single molecule images or to fit them, the width parameter of the Gaussian pixelated profile was calculated as $\sigma = 1.323\lambda/2\pi n_a = 84.225 \ nm$.

noise. The location of the single molecule was estimated from each image by fitting the pixelated profiles using both the nonlinear least squares and maximum likelihood algorithms, again assuming the absence of any model misspecifications. The type of profile used to fit the data matched the type of profile used to generate the data, i.e. Airy pixelated profiles were used to
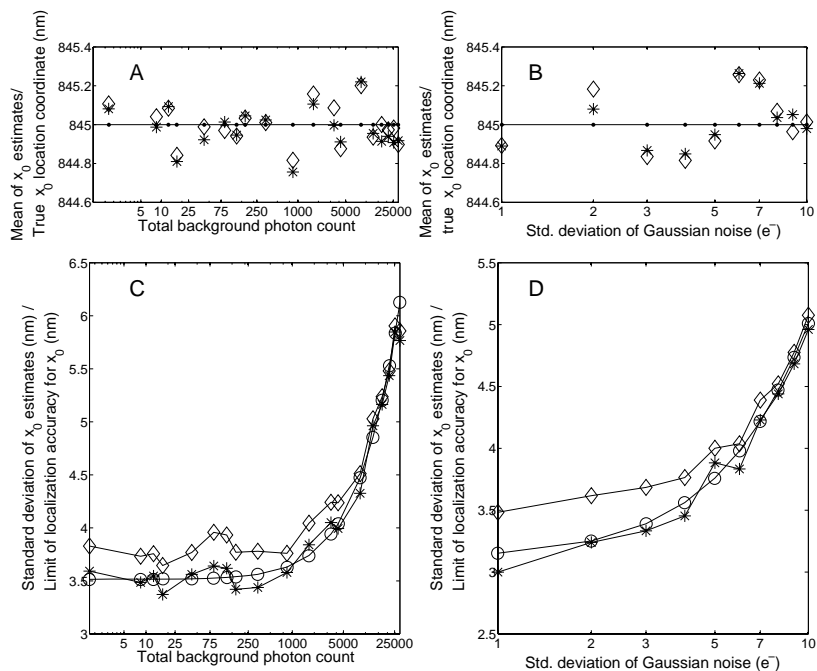
Fig. 3. Comparison of standard deviation of estimates from the nonlinear least squares and maximum likelihood estimators in the presence of extraneous Poisson and Gaussian noise sources and the absence of model mismatches. Panels $A(C)$ and $B(D)$ shows the mean(standard deviations) of the x_0 coordinate estimates from the nonlinear least squares ($\diamond$) and maximum likelihood ($*$) estimators. ($\bullet$) indicates the true x_0 coordinate value, and ($\circ$) the PLAM or limit of the localization accuracy of x_0. In panels A and C, 1000 images of a stationary single molecule were generated using the Airy pixelated profile and readout noise with standard deviation of $4e^-$ for each value of the total background photon count. In panels B and D, 1000 images of a stationary single molecule were generated using the Airy pixelated profile and fixed background of 2 photons/pixel/s for each value of the standard deviation of Gaussian noise. The mean of the Gaussian noise component was set to zero in all cases. Location coordinates (x_0, y_0) were estimated from each image using both estimators. Values for the width, photon detection rate, and background parameters, were fixed to the values used to generate the images. The following numerical values were used when simulating the single molecule images. Pixel size: $13\ \mu m \times 13\ \mu m$, pixel array size: 13×13, expected number of photons from the single molecule at the detector plane: 1000 photons, magnification $M = 100$, wavelength $\lambda = 520\ nm$, numerical aperture $n_a = 1.3$. The single molecule image was centered within the pixel array.

fit data generated using Airy pixelated profiles, and likewise for Gaussian pixelated profiles. Both algorithms are again able to recover the true location of the single molecule on average [Fig. 2(a) and 2(b)]. Examining the standard deviations of the location estimates from both algorithms, we see that for low noise levels, relative to the total photon count from the single molecule, the maximum likelihood estimator has a better accuracy, as evidenced by the smaller standard deviations, than the nonlinear least squares estimator. However, for high noise levels, the standard deviations of estimates obtained from both algorithms matched the PLAM. To further study the effect of Gaussian noise on the performance of these two algorithms, we followed two approaches. First we added a Gaussian noise component with standard deviation
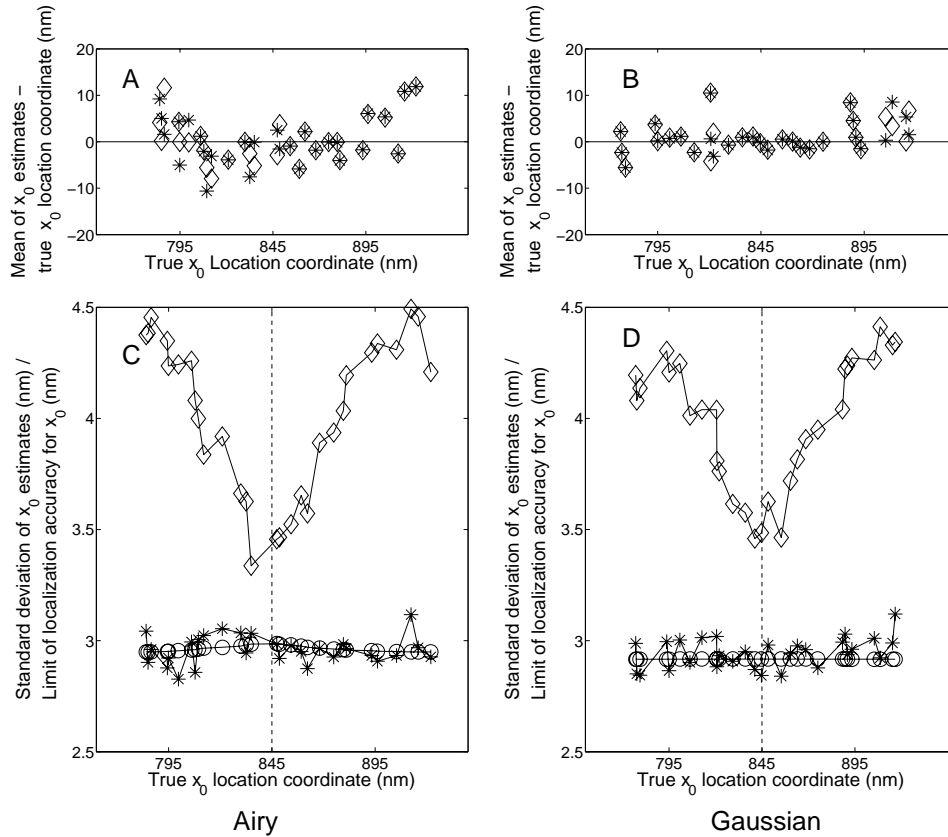
Fig. 4. Comparison of standard deviation of estimates from the nonlinear least squares and maximum likelihood estimators when the point source is away from the center of the image array. Panels A and B show the difference between the mean of the x_0 coordinate estimates from both algorithms and the true x_0 coordinate value. Panels C and D show the standard deviations of the x_0 coordinate estimates from both estimators. ($\diamond$) show the results from the nonlinear least squares estimator, ($*$) the results from the maximum likelihood estimator, and ($\circ$) the PLAM or limit of the localization accuracy of x_0. The dashed line indicates the center of the image array in absolute coordinates. For each position of the point source, 1000 images of a stationary single molecule were generated using either the Airy (panels A and C) or Gaussian (panels B and D) pixelated profile. Location coordinates (x_0, y_0) were estimated from each image using both estimators. Values for the width, photon detection rate, and background parameters, were fixed to the values used to generate the images. The following numerical values were used when simulating the single molecule images. Pixel size: 13 $\mu m \times$ 13 μm, pixel array size: 13×13, expected number of photons from the single molecule at the detector plane: 1000 photons, magnification $M = 100$. When the Airy pixelated profile was used to generate the single molecule images or to fit them, wavelength $\lambda = 520$ nm, and numerical aperture $n_a = 1.3$. When the Gaussian pixelated profile was used to generate the single molecule images or to fit them, the width parameter of the Gaussian pixelated profile was calculated as $\sigma = 1.323\lambda/2\pi n_a = 84.225$ nm. The x_0 coordinate of the center of the pixel array in the object space is 845nm.

of $4e^-$ and zero mean to the data used in the calculations above and again estimated the location of the single molecule using both the nonlinear least squares and the maximum likelihood

(using the Gaussian noise model) estimators. Second, for a fixed background photon detection rate of 2 photons/pixel/second (total background photon count of 338 photons), we generated images of a single molecule with varying levels of Gaussian noise and used both algorithms to again estimate the location of the single molecule. Both algorithms are again able to recover the true location of the single molecule on average in both scenarios [Fig. 3(a) and 3(b)]. The results from the first approach [Fig. 3(c)] with a constant Gaussian noise of standard deviation $4e^-$ show that for low Gaussian noise levels the performance of the two estimators is almost identical to the case when there is no Gaussian noise. The results from the second approach [Fig. 3(d)] show that increasing the Gaussian noise component also causes the standard deviations of the two algorithms to converge to the PLAM. Thus, neither algorithm provides a clear accuracy advantage when analyzing data containing a large noise component. However, at low noise and signal levels, there is an appreciable difference in the accuracy of estimates from the two algorithms.

3.4. The maximum likelihood estimator is less sensitive to the relative location of the point source within the pixel

In all the estimations performed so far, the point source was placed at the center of the detector pixel array to generate the data. To study the performance of the two algorithms when the point source is positioned away from the center of the image array, we generated sets of one thousand images of a single molecule whose center was changed in each set of images in increments of $5nm$ in the object space starting from the left edge of the center pixel and moving toward the right edge of the center pixel of the image array. The single molecule images were generated using both pixelated Airy and Gaussian profiles. The location of the single molecule was estimated from each image using both the nonlinear least squares and maximum likelihood estimators, again assuming the absence of model misspecifications. Both algorithms are again able to recover the true location of the single molecule on average. However, examining the standard deviations of the estimates from both algorithms, we see that while the maximum likelihood algorithm attains the PLAM irrespective of the position of the point source, the accuracy of the nonlinear least squares estimator deteriorates the further away from the center of the pixel the point source is located. Thus the maximum likelihood algorithm again achieves the best possible accuracy in this scenario.

3.5. The maximum likelihood estimator is less sensitive to misspecifications of the width parameter

In all the estimations performed so far, the values for the fixed parameters (width, photon detection rate, etc.) were accurately specified for the profile being fit, i.e., they were set to the values used to generate the images. In many practical situations, values for parameters such as the width parameter of the estimated profile can be misspecified because of various factors. We have previously shown that depth discrimination is very poor close to the plane of focus (2). Therefore, when visually focusing on a sample, there is a likelihood of being away from the true plane of focus by a few hundred nanometers. In our experience, the width parameter of the PSF can deviate by up to 50% from the theoretical width parameter, as specified by the experimental and imaging conditions, if the image being fit is out-of-focus by a few hundred nanometers. We performed an analysis using simulations to verify this. Using the Born-Wolf PSF model, for various levels of defocus [22, 25], we simulated 13×13 pixel image profiles of a single molecule emitting light of wavelength $\lambda = 520nm$, imaged through an objective of numerical aperture $n_a = 1.3$ and magnification $M = 100$, acquired on a camera with pixel size $13\mu m \times 13\mu m$. We then fit these image profiles using an Airy PSF with the same experimental and imaging conditions while floating the location coordinates and the width parameter. The theoretical value

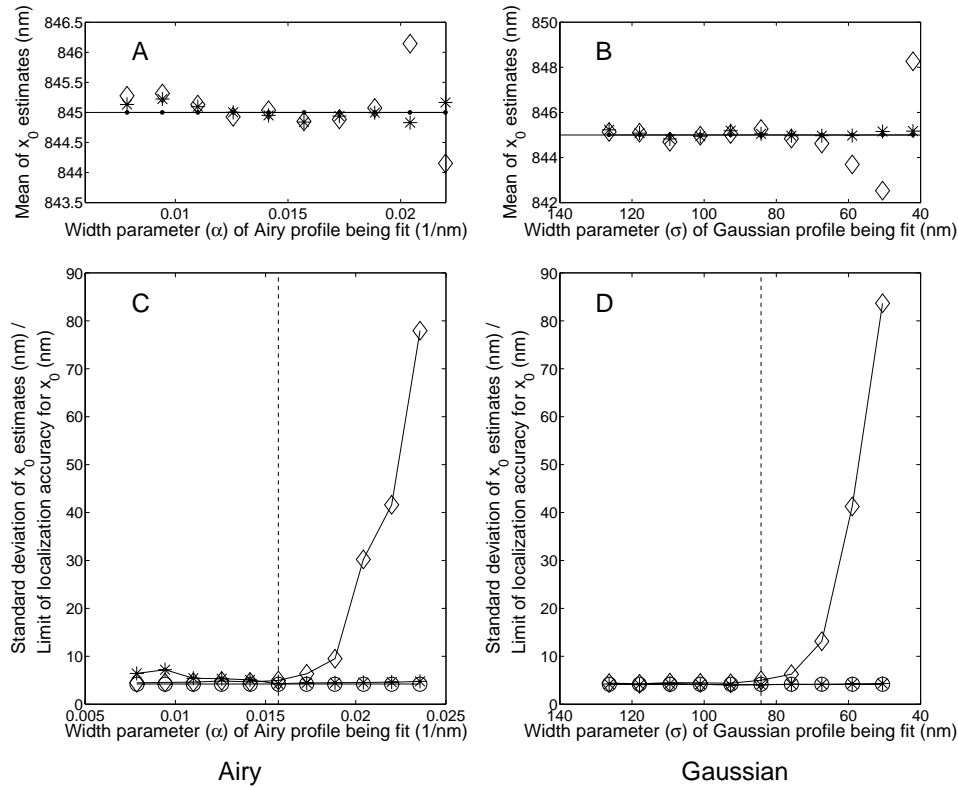Airy   Gaussian

Fig. 5. Comparison of the accuracies of the nonlinear least squares and maximum likelihood estimators when the data is generated and fit with the same type of profile but the width parameter is misspecified. Panels *A*(*B*) and *C*(*D*) show the results when the Airy (Gaussian) pixelated profile was used both to generate the images of the single molecule and to fit them. Panels *A*(*C*) and *B*(*D*) show the mean (standard deviations) of the x_0 coordinate estimates from both the nonlinear least squares ($\diamond$) and the maximum likelihood ($*$) estimators. ($\bullet$) indicates the true x_0 coordinate value. ($\circ$) indicates the PLAM or limit of the localization accuracy of x_0. The dashed line indicates the correct value for the width parameter. Values for the photon detection rate and background parameters were fixed to their true values, i.e., the values used to generate the images. In both cases, sets of 1000 images of a stationary single molecule were generated. The location coordinates (x_0, y_0) were estimated from each image using both algorithms. The width parameter of the profile being fit was misspecified by various amounts between sets of images. The following numerical values were used when simulating the single molecule images. Pixel size: 13 $\mu m \times 13$ μm, pixel array size: 13×13, magnification $M = 100$. The single molecule image was centered within the pixel array. The mean of the background noise component was set to zero. When the Airy pixelated profile was used to generate the single molecule images, wavelength $\lambda = 520$ nm, and numerical aperture $n_a = 1.3$, resulting in $\alpha = 0.0157$ nm^{-1}. When the Gaussian pixelated profile was used to generate the single molecule images, the width parameter of the Gaussian pixelated profile was calculated as $\sigma = 1.323\lambda/2\pi n_a = 84.225$ nm.

for the width parameter, given by $2\pi n_a/\lambda$, is $15.7\mu m^{-1}$. However, at just $200nm$ defocus, the value for the width parameter of the Airy profile was estimated at $6.53\mu m^{-1}$, approximately half of the true value.

Differences between the width parameter used when fitting profiles and the true width param-

eter can also occur if the width parameter is determined solely by theoretical considerations, as properties of the optical components can deviate from what is specified by the manufacturer [26]. In the example given above, the true value of the width parameter value, assuming that the operational numerical aperture of the objective matches the design numerical aperture, was $15.7 \mu m^{-1}$. However, if the operational numerical aperture was 1.2, the true value of the width parameter would be $14.5 \mu m^{-1}$. Thus, variations in the experimental properties of optical components along with the possibility of the sample being out of focus can lead to large deviations in the width parameter.

Consequently, it is important to understand how the misspecification of the width parameter affects the performance of the estimation algorithms. Here we investigated the effects of a misspecification of the width parameter. To study these effects, we simulated sets of one thousand images of a stationary single molecule using both Airy and Gaussian pixelated profiles. The single molecule location was estimated from each image, using both the maximum likelihood and nonlinear least squares estimators, but the width parameter of the profile being fit was deliberately misspecified by a certain amount for each set of images. The type of profile used to fit each image matched the type of profile used to generate that image, i.e., images generated using pixelated Airy profiles were fit using pixelated Airy profiles, and images generated using pixelated Gaussian profiles were fit using pixelated Gaussian profiles.

Examining the location estimates, the results show that both the maximum likelihood and least squares estimators are, on average, able to recover the true location of the single molecule irrespective of the amount by which the width of the profile being fit is misspecified [Fig. 5(a) and 5(b)]. Examining the standard deviations of the estimates calculated independently from each set of images for each algorithm [Fig. 5(c) and 5(d)], the standard deviations of the estimates obtained using the maximum likelihood estimator are again close to the limit of the localization accuracy, irrespective of the amount by which the width parameter of the profile being fit is misspecified, thus showing that the accuracy of the maximum likelihood estimator is relatively unaffected by the misspecification of the width parameter.

On the other hand, the accuracy of the nonlinear least squares estimator depends on whether the width parameter value causes the profile being fit to be wider or narrower than the profile of the single molecule image. If the width of the profile being fit is wider than the profile of the single molecule, then the standard deviations of the estimates from the nonlinear least squares estimator are also close to the limit of the localization accuracy. However, setting the width parameter of the profile being fit to values that cause it to be narrower than the profile of the single molecule leads to larger standard deviations in the estimates from the nonlinear least squares estimator. Thus the nonlinear least squares estimator appears to be more sensitive to misspecification of model parameters while the maximum likelihood estimator is more robust in the presence of model misspecifications.

3.6. Model mismatch has different effects on the performance of the two algorithms

Gaussian profiles are often used to fit data that might be more accurately modeled by Airy profiles resulting in a mismatch between the model by which the data is generated and the model used to fit the data. A key step in approximating Airy profiles using Gaussians is determining a value for the width parameter of the Gaussian that produces a profile that best approximates an Airy profile. For this purpose, we fit the Gaussian photon distribution profile function [see Appendix A, Eq. (5)] to the Airy photon distribution profile function [see Appendix A, Eq. (4)] for various values of α, using criteria based on the nonlinear least squares and the maximum likelihood function [see Appendix B, Eqs. (8) and (6)]. Both criteria produced similar results. From them, we determined that for a Gaussian profile that best approximates an Airy profile with width parameter α, the associated Gaussian width parameter σ can be calculated by the
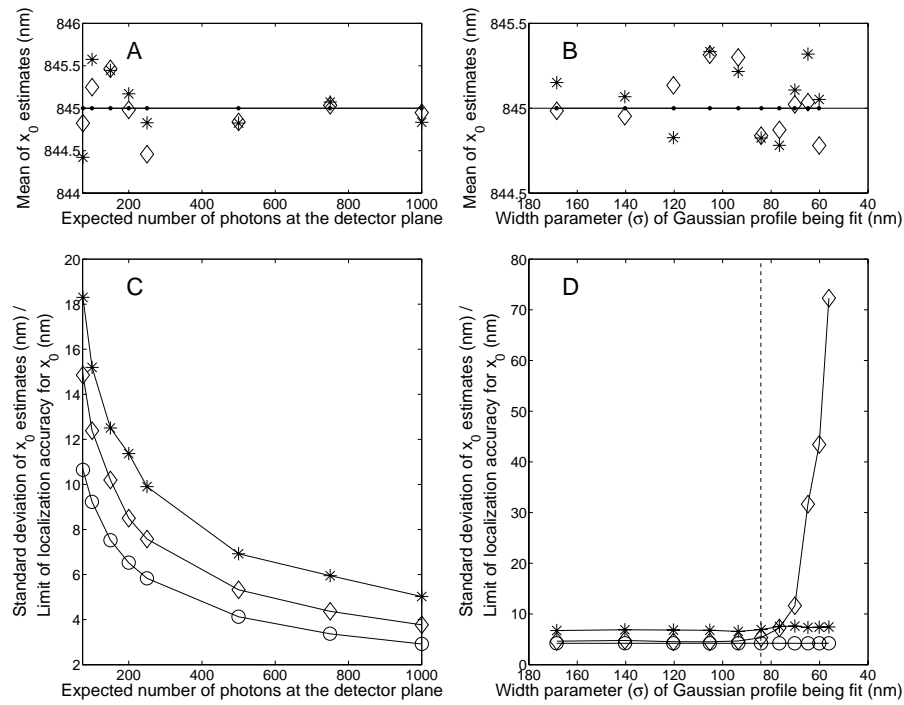
Fig. 6. Comparison of the accuracies of the nonlinear least squares and maximum likelihood estimators when Gaussian pixelated profiles are used to fit data generated using Airy pixelated profiles. Panels $A(C)$ and $B(D)$ show the mean (standard deviation) of the x_0 coordinate estimates from both the nonlinear least squares ($\diamond$) and the maximum likelihood ($*$) estimators. ($\bullet$) indicates the true x_0 coordinate value. ($\circ$) indicates the PLAM or limit of the localization accuracy of x_0. In panels A and C, sets of 1000 images were generated for each value of the expected photon count at the detector plane. In panels B and D, sets of 1000 images were generated with the expected photon count at the detector plane set to 500 photons. In both cases, the location coordinates (x_0, y_0) were estimated from each image using both the nonlinear least squares and maximum likelihood estimators. Values for the photon detection rate and background parameters were fixed to their true values, i.e., the values used to generate the images. In panels A and C the value for the width parameter was also fixed to the true value. All images were generated using Airy pixelated profiles and fit using Gaussian pixelated profiles. In panels B and D, the width parameter of the Gaussian profile was misspecified by varying amounts between sets of images. The dashed line indicates the best approximate for the Gaussian width parameter. The following numerical values were used when simulating the single molecule images. Pixel size: $13 \, \mu m \times 13 \, \mu m$, pixel array size: 13×13, magnification $M = 100$. The single molecule image was centered within the pixel array. The mean of the background noise component is set to zero. For the Airy pixelated profile used to generate the single molecule images, wavelength $\lambda = 520 \, nm$, and numerical aperture $n_a = 1.3$, resulting in $\alpha = 0.0157 \, nm^{-1}$. For the Gaussian pixelated profile used to fit the images (in the absence of misspecifications), the width parameter was calculated as $\sigma = 1.323\lambda/2\pi n_a = 84.225 \, nm$.

approximation $\sigma \approx 1.323/\alpha$ which is in agreement with the results published previously [13, 27].

To study the effect of the model mismatch considered here on the accuracy of the two esti-

mators, we investigated two scenarios. In the first scenario, we simulated sets of one thousand images of a stationary single molecule for various values of the expected number of photons at the detector plane using Airy pixelated profiles, and fit them with Gaussian pixelated profiles of specified width $\sigma = 1.323/\alpha$ to estimate the single molecule location coordinates. In this scenario, the values for the other parameters, e.g. photon detection rate, were fixed to the values used to generate the single molecule images. In the second scenario, we simulated multiple sets of one thousand images using Airy pixelated profiles for a fixed value of the expected number of photons at the detector plane, and again fit them with Gaussian pixelated profiles to estimate the single molecule location coordinates. However, for each set of images, we deliberately misspecified the width parameter of the Gaussian profile being fit by a certain amount. The values for the photon detection rate and background parameters were again fixed to the values used to generate the single molecule images. Images were fit using both the nonlinear least squares and maximum likelihood estimators.

From the results, we see that in both scenarios both estimators are, on average, able to recover the single molecule location despite the model misspecification [Fig. 6(a) and 6(b)]. Calculating the standard deviations of the location estimates independently from each set of images for each algorithm, we see that the standard deviations from neither algorithm matches the PLAM in either scenario. In the first scenario, we see that with this specific model mismatch, the accuracy of the nonlinear least squares estimator is better than that of the maximum likelihood estimator as evidenced by the smaller standard deviations [Fig. 6(c)]. In the second scenario, we see that the the maximum likelihood estimator is no longer able to attain the PLAM. However, we do observe that the standard deviations of the estimates from the maximum likelihood estimator are relatively constant, irrespective of the width of the profile being fit [Fig. 6(c)]. This again shows that the maximum likelihood estimator is robust in the presence of model misspecifications. On the other hand, we find that the standard deviations of the estimates from the nonlinear least squares estimator again depend on whether the width parameter value causes the profile being fit to be wider or narrower than the profile of the single molecule image. If the width of the profile being fit is wider than the profile of the single molecule, then the standard deviations of the estimates from the nonlinear least squares estimator are smaller than those from the maximum likelihood estimator. However, the narrower the width of the profile being fit compared to the profile of the single molecule, the larger are the standard deviations of the estimates from the nonlinear least squares estimator, indicating a deterioration in accuracy [Fig. 6(d)]. This again shows the sensitivity of the nonlinear least squares estimator to model misspecifications.

3.7. Two analytical approaches predict different localization accuracies

Analytical expressions for the accuracy with which the location parameters, and other parameters, can be estimated are important tools for the design of single molecule experiments and for the evaluation of location estimation techniques. In Fig. 1(b) we showed that the theoretically optimal standard deviations as expressed by the PLAM can in fact be attained by the maximum likelihood estimator [13, 14]. Another expression for the prediction of the standard deviation of a location estimator has previously been published by Thomspon et. al. [17]. Since both approaches have been used in previous studies to predict localization accuracies, we wanted to compare the two analytical approaches to evaluating the performance of location estimators.

The formula by Thompson et. al. for an estimate of the localization error for a given set of imaging and experimental conditions is given by

$$\langle (\Delta x)^2 \rangle = \frac{s^2}{N} + \frac{a^2/12}{N} + \frac{8\pi s^4 b^2}{a^2 N^2},\tag{1}$$

where Δx denotes the standard deviation of the single molecule location estimates, s denotes the standard deviation of the Gaussian image profile that is assumed to describe the image of the
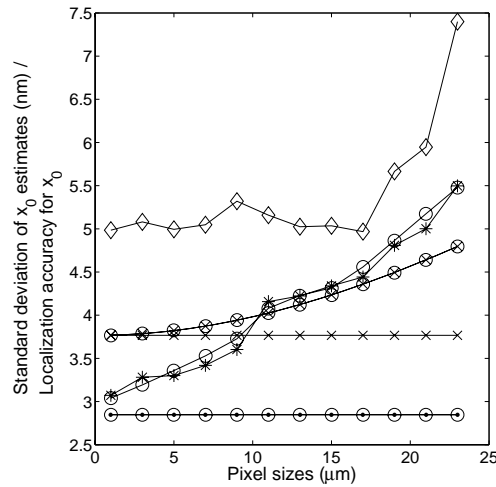
Fig. 7. Comparison of the FLAM and PLAM with estimates of the localization error according to Thompson et. al. [17]. The PLAM for the Airy pixelated profile ($\circ$) and estimates of the localization error given by Eq. 1 [17] ($\otimes$) are plotted as a function of pixel size. For comparison, the FLAM ($\odot$) and the corresponding localization error estimate by Thompson et. al. for the ideal case ($\times$), calculated as $s/\sqrt{N}$, are also plotted. For each pixel size, a set of 1000 single molecule images were generated using the Airy pixelated profile. The standard deviation, calculated independently for each pixel size, of the single molecule location coordinate estimates from each image obtained using both the nonlinear least squares ($\diamond$) and maximum likelihood ($*$) estimators are also shown. The numerical values used in the calculations are as follows. Pixel array size: closest match to $169\mu m \times 169\mu m$ total detector area, magnification $M = 100$, expected number of photons from the single molecule: 500 photons. The single molecule image was centered within the detector area. The mean of the background noise component is set to zero. For the Airy profile, wavelength $\lambda = 520$ nm, and numerical aperture $n_a = 1.3$. For the calculations of the localization error based on Eq. 1, s is calculated as $1.323\lambda/2\pi n_a = 84.225$ nm, $N = 500$ (expected number of photons from the single molecule), and a is obtained by dividing the pixel size by the magnification.

single molecule, a the length of a pixel of the detector in the object space (pixels are assumed to be square), N the total photon count from the single molecule, and b the standard deviation of the background [17].

If the image profile is Gaussian, and pixelation and other extraneous noise sources are not accounted for, then in fact the two approaches lead to the same result, i.e., $\sigma/\sqrt{N}$, where N is the expected number of collected photons and σ denotes the parameter that describes the 'standard deviation' of the Gaussian image profile [14]. In practice, when the image of the single molecule is not necessarily a Gaussian function, it is suggested that the experimental profile should be fitted with a Gaussian profile and that the corresponding σ parameter should be used in the analytical expression above [17, 28]. We therefore investigated how well the two approaches for predicting the localization accuracy perform when the image of the single molecule is best described by an Airy profile.

In the absence of deteriorating influences such as pixelation and extraneous noise sources, the FLAM for a single molecule with an Airy image profile is given by $1/(\alpha\sqrt{N})$, with $\alpha = 2\pi n_a/\lambda$ and N the expected number of photons at the detector plane as described above. As

discussed earlier, for a Gaussian profile that approximates an Airy profile, the standard deviation parameter σ can be approximated by $\sigma \approx 1.323/\alpha$. Substituting in Eq. 1 (while disregarding pixelation and noise sources), we obtain $\sigma/\sqrt{N} = 1.323/(\alpha\sqrt{N})$ for the predicted standard deviation of the location estimates. This implies that there is approximately a 30% difference in the accuracy predicted by the two analytical approaches for this particular scenario.

In Fig. 7 the localization accuracies predicted by both analytical approaches are compared when pixelation is taken into account. We see that the standard deviations predicted by the approach based on Eq. (1) converges to the FLAM for a Gaussian image profile as the pixel size is reduced, while the PLAM for an Airy image profile converges to the FLAM for an Airy image profile. We also observe that for small pixel sizes the expression based on Eq. (1) predicts values for the standard deviation of the estimator that are higher than the PLAM for an Airy profile. For larger pixel sizes, the expression based on Eq. (1) predicts standard deviations that are smaller than the PLAM.

The difference in the results predicted by these two analytical approaches was further verified through simulations. We generated sets of one thousand images of a stationary single molecule for the various pixel sizes in Fig. 7 using pixelated Airy profiles. The location of the single molecule was estimated from each image by fitting pixelated Airy profiles using both the nonlinear least squares and maximum likelihood estimators. The values for the other parameters, e.g. width and photon detection rate, were fixed to the values used to generate the single molecule images. Calculating the standard deviations of the estimates from each algorithm independently for each of the pixel sizes considered, we see that the standard deviations of the estimates from the maximum likelihood estimator match the PLAM. The standard deviations of the estimates obtained using the nonlinear least squares estimator match neither the PLAM nor the expression by Thompson et. al. [17]. These observations may help to explain reports in the literature that the standard deviations from the nonlinear least squares estimator are not always adequately explained by Eq. (1). In contrast, they suggest that lower standard deviations can be achieved as expressed by the PLAM and as attained by the maximum likelihood estimator.

4. Discussion

We have examined two estimators, the nonlinear least squares estimator and the maximum likelihood estimator, with respect to their performance for estimating the location of single molecules or other point sources. Amongst a selection of estimators that did not contain the maximum likelihood estimator, the nonlinear least squares estimator had been identified as the most accurate in a comparative study [11]. This was done in the context of fitting a Gaussian profile to data generated with an Airy model. The maximum likelihood estimator, however, is a classical estimator [12, 29] that has also been investigated in the context of single molecule localization [13, 14]. It was therefore important to extensively compare these two estimation approaches.

The first criterion for evaluation is whether or not an estimator recovers the correct parameter, in this case, the coordinates of the location of the single molecule. Considering the stochastic nature of the data, it cannot be expected that the correct parameter can be exactly recovered. However, it is shown here with simulations that both estimators, on average, can retrieve the location of single molecules without any obvious bias. The performance of estimators for the localization of single molecules is typically evaluated based on the standard deviation or variance of the resulting estimates [13, 17]. This is consistent with the desire that an estimator, which on average produces the correct results, does so with a small standard deviation. Our results show that in the absence of model misspecifications, the maximum likelihood estimator consistently produces estimates with lower standard deviation than the nonlinear least squares estimator.

In any practical situation, the effect of extraneous noise (background autofluorescence, camera readout noise, etc) on the performance of estimation algorithms is an important consideration. We have shown that as the extraneous noise component in the data increases, irrespective of whether the noise is Gaussian (camera noise) or Poisson (background autofluorescence) in nature or a combination of both, the standard deviations of the estimates from the nonlinear least squares algorithm become comparable to those from the maximum likelihood estimator and converge to the PLAM. This behavior of the estimators is not completely surprising as similar results has been observed in deconvolution studies [15, 16]. We know from the statistical literature that the nonlinear least squares algorithm is better suited for data with a Gaussian distribution while the maximum likelihood algorithm is better suited for data with a Poisson distribution. Consider the scenario when the data contains only signal from the point source and Poisson noise. At high signal levels, i.e. when there is a large Poisson noise component in this scenario and consequently a non-trivial signal measurable at every pixel, we know that data with a Poisson distribution can be approximated well by a Gaussian distribution. Since the nonlinear least squares algorithm is suited for data that exhibits Gaussian behavior, it is not surprising that for data with a large Poisson noise component the nonlinear least squares algorithm produces results with accuracies comparable to those from the maximum likelihood estimator. In the case when there is a large readout noise component, the Gaussian nature of the readout noise dominates the nature of the overall signal. Therefore, it can again be expected that the performance of the nonlinear least squares algorithm would be comparable to that of the maximum likelihood algorithm. However, in the realm of quantum limited data, i.e. data where the overall signal level is low at a significant number of pixels in the image, as is the case with single molecule data, the Poisson nature of the data dominates and the maximum likelihood estimator is better suited in this situation. Consequently, at low extraneous noise levels, the maximum likelihood algorithm produces more accurate estimates.

From a practical point of view the question of robustness is of great importance. In a concrete experimental situation the acquired data may not precisely match the theoretical assumptions. For example, the image profile might deviate from the assumptions due to the presence of unmodeled optical effects or due to the fact that some of the parameters that specify the profile might not be accurately known. The robustness question now relates to what extent the behavior of the estimator changes if the matching image profile is not perfectly modeled. Importantly, for the model mismatches that were considered here, both estimators, on average, recover the correct location parameters. For the nonlinear least squares estimator the standard deviations can be significantly larger than in the ideal case. If the model is misspecified this estimator can even produce better results than when the model is correctly specified, as is the case when the specified width parameter produces a profile that is considerably wider than the single molecule image profile. In contrast the performance of the maximum likelihood estimator does not depend very significantly on the nature and size of the misspecification.

It should, however, be emphasized that it cannot be expected that the estimators behave properly under any distortions. For example, if the actual data and the fitted profiles no longer exhibit radial symmetry, different results should be expected. If very precise results are desired, our analysis suggests that great care has to be taken that the model that is used for fitting is highly accurate. In general, if a certain type of modeling uncertainty is expected and very highly accurate results are required, it might be important to carry out an analysis of the type performed here for the specific modeling uncertainty. In a number of scenarios it was shown that the maximum likelihood estimator achieves the best results especially in the ideal case, i.e., when no model mismatch is present. Importantly, in this case, it attains the best possible performance as predicted by the PLAM [13]. In addition, it does not exhibit the potentially very large performance deteriorations that are seen with the nonlinear least squares estimator. Based

on the analysis carried out here the maximum likelihood estimator might therefore be a better algorithm choice.

The objective function of the maximum likelihood estimator is chosen based on the nature of the underlying data. In this study, we have used two objective functions with the maximum likelihood estimator, one for data with only Poisson characteristics, and a computationally more complex one for Poisson distributed data with additive Gaussian noise. The objective function of the nonlinear least squares estimator, however, remains the same irrespective of the nature of the data for which it is being used. The execution speed of the maximum likelihood estimator using the objective function for data with only Poisson characteristics is comparable to the execution speed of the nonlinear least squares estimator. However, the maximum likelihood estimator using the objective function for Poisson distributed data with additive Gaussian noise is computationally more demanding than the nonlinear least squares estimator because of the complexity of the objective function. We have shown that for a number of important scenarios, the maximum likelihood estimator using the more complex objective function produces more accurate results than the nonlinear least squares estimator. In such scenarios, it is for the microscopist to decide whether the increased computational burden of using the maximum likelihood estimator with the more complex objective function is justified by the increased accuracy.

The comparison between the two analytical expressions for the localization accuracy showed that differences can occur in the predicted levels of accuracy. The expressions based on the Fisher information matrix had been derived using a rigorous statistical framework. In contrast, the approach in [17] relied on ad hoc derivations. Both approaches lead to identical results when the image profile is Gaussian and no pixelation is assumed. The presence of pixelation will also, in many cases, not lead to significant differences. Both approaches can, however, lead to significantly different results when the image profile is not Gaussian. The Fisher information matrix based approach can be used for any point spread function model. In the approach by Thompson et. al. [17], an approximation is typically carried out, whereby the best Gaussian approximate to the actual point spread function profile is determined. Assuming that the actual point spread function is an Airy profile, the two approaches were compared. It was seen that in this situation the two approaches typically no longer predict the same accuracy with which the location of the single molecule can be estimated. The lowest standard deviations are predicted by the PLAM, i.e., by the Fisher information based approach and are also attained by the maximum likelihood estimator. Thus the difference between the localization accuracy predicted by the approach utilizing the Fisher information matrix and the approach by Thompson et. al. arise from the assumptions made about the underlying data model. Since these two approaches produce different results, a microscopist should take this difference into consideration when comparing the experimentally observed localization error with the theoretically predicted one. Many factors can lead to differences between the experimentally observed and theoretically determined localization accuracies, e.g. inadequate correction for stage drift, optical aberrations. Our comparison of the two approaches to predict the localization error suggests that the approach used to determine the localization error can also contribute to these differences. Our results show that the standard deviations of estimates from the maximum likelihood estimator are consistent with the PLAM in the absence of model mismatches. Therefore, using the Fisher information matrix based approach to calculate localization accuracies can help minimize differences between the experimentally observed and theoretically determined localization errors.

In the present study, we are primarily interested in the performance of the nonlinear least squares and maximum likelihood algorithms with respect to localization accuracy. Therefore, in most cases we have estimated only the location coordinates and fixed various other parameters (e.g. width, photon detection rate, background). In our experience, floating some of these parameters in addition to the location parameters can cause significant problems with the esti-

mation. For example, we have observed that floating the background parameter along with the width parameter can lead to significant deterioration of the results. Estimating fewer parameters also has advantages with regard to computational efficiency. This requires more parameters to be independently specified. Previously, we have theoretically shown, and numerically confirmed here, that the estimation of the photon detection rate is uncorrelated with the estimation of the location of the single molecule or point source [14]. This means that the photon detection rate and location parameters can be estimated independently without any change in the accuracy with which either parameter is estimated. We have here also explored how errors in specifying the width parameter affects the accuracy of location estimates. From the results, we see that misspecifications in the width parameter in one direction do not affect the accuracy of the estimates from either algorithm in some scenarios. This implies that care must be taken to make sure that if misspecifications occur they are such that they do not produce severely deteriorated estimates.

The computations required to estimate the location of a single molecule and to evaluate the limit of the localization accuracy can sometimes be fairly complex. To make these methodologies available to users who do not wish to write their own code, we have developed software packages to accomplish these tasks. The EstimationTool [23] allows users to perform various estimation tasks including single molecule location estimation and resolution/distance measurements in 2D and 3D. It offers different choices of estimation models (Airy, Gaussian, Born-Wolf 3D point spread function model [25]), the ability to use either the nonlinear least squares or the maximum likelihood estimator, and supports various models for extraneous noise sources (Poisson only, Poisson and Gaussian noise). The tool also provides access to advanced calculation parameters, for example parameters involved in performing the numerical integrations.The EstimationTool also allows results of estimations to be visualized and exported for further analysis. All calculations of the limit of the localization accuracy can be performed with the FandPLimitTool [24]. Using the FandPLimitTool both localization and resolution measures can be calculated in 2D and 3D for various estimation models. User-friendly graphical user interfaces are available for both packages [23, 24].

Appendix A: Simulating the single molecule images

To compare the accuracies of the nonlinear least squares and maximum likelihood estimators, we simulated images of a stationary single molecule as acquired by a camera and estimated the location of the single molecule using both estimators. When simulating the single molecule images, the camera is modeled as a set of pixels denoted by $\{C_1, \ldots, C_K\}$. The signal acquired at the kth pixel is modeled as

$$\mathscr{I}_{\theta,k} := S_{\theta,k} + B_k + W_k, \tag{2}$$

where θ denotes the vector of parameters to be estimated, $S_{\theta,k}$ is a Poisson random variable denoting the signal from the single molecule detected at the kth pixel, and B_k is a Poisson random variable denoting the signal from the background component with mean $b\Delta t$ where b is the rate at which photons from the background are being detected and Δt is the time interval over which the image is acquired, and W_k is a Gaussian random variable denoting the camera noise with mean $\eta_{w,k}$ and standard deviation denoted by $\sigma_{w,k}$.

The mean of $S_{\theta,k}$ is given by

$$\mu_{\theta,k} := (A \cdot \Delta t) \int_{C_k} f_\theta(r) \, dr, \tag{3}$$

where A denotes the rate at which photons from the single molecule are being detected, Δt denotes the time interval over which the image is acquired, C_k denotes the area of the kth pixel,

and f_θ denotes the photon distribution profile, a probability distribution function which gives the probability of a photon being detected at a specified location [13].

For the simulation of an Airy profile, f_θ is defined as

$$f_\theta(r) := \frac{\left[J_1\left(\frac{\alpha}{M}||r-r_0||\right)\right]^2}{\pi||r-r_0||^2},\tag{4}$$

where $||r-r_0|| = \sqrt{(x-Mx_0)^2 + (y-My_0)^2}$, M denotes the magnification of the optical system, $\theta := (x_0, y_0)$ is the location of the single molecule in the object space, and $1/\alpha$ denotes the width of the profile with $\alpha := 2\pi n_a/\lambda$, where n_a is the numerical aperture of the system, and λ is the wavelength of the light emitted by the single molecule.

For the simulation of a Gaussian profile, f_θ is defined as

$$f_\theta(r) := \frac{1}{2\pi M^2 \sigma^2} \cdot e^{\left(-\frac{(x-Mx_0)^2}{2(M\sigma)^2} - \frac{(y-My_0)^2}{2(M\sigma)^2}\right)},\tag{5}$$

where σ denotes the width of the Gaussian profile.

To simulate a single molecule image, the mean pixel intensity was calculated using Eq. (3) for each of the K pixels, leading to a set $\{\mu_{\theta,1} + b\Delta t, \ldots, \mu_{\theta,K} + b\Delta t\}$ of mean pixel intensity values, which is referred to as the pixelated profile. For each image, a Poisson realization of the pixelated profile was generated. To simulate additive Gaussian noise, independent realizations of the Gaussian random variable W_k in Eq. (2) for a fixed value of $\sigma_{w,k}$ are added to each pixel in the Poisson realization of the pixelated profile. In all calculations, a set of images refers to a set of one thousand simulated images. All calculations involved in generating the single molecule images were performed in MATLAB (The MathWorks, Natick, MA). The integration routines used to integrate the photon distribution profile function over a pixel were custom developed in MATLAB. Poisson realizations were generated using the Poisson random number generator provided with MATLAB's Statistics Toolbox.

Appendix B: Fitting single molecule images

Both the maximum likelihood and nonlinear least squares algorithms were used to fit the single molecule images and estimate the location of the single molecule. Both algorithms fit the images using the pixelated profiles calculated from Eq. (3) using either the Airy or Gaussian photon distribution profile functions given by Eq. (4) or Eq. (5).

The nonlinear least squares estimate was obtained by minimizing the function

$$S = \sum_{k=1}^{K} \left[z_k - \left(\mu_{\theta,k} + b\Delta t\right)\right]^2,\tag{6}$$

where $z_1, \ldots, z_K$ denotes the single molecule image data, and $\mu_{\theta,k}$ denotes the mean pixel intensity as given by Eq. (3). MATLAB's lsqnonlin function from the Optimization Toolbox was used to perform the nonlinear least squares minimization.

The maximum likelihood estimate for data incorporating Gaussian noise was obtained by minimizing the negative of the function

$$\mathscr{L}(\theta|z_1, \ldots, z_K) := \sum_{k=1}^{K} \ln \left(\frac{1}{\sqrt{2\pi}\sigma_{w,k}} \cdot \sum_{l=0}^{\infty} \left(\frac{[\mu_{\theta,k} + b\Delta t]^l e^{-(\mu_{\theta,k} + b\Delta t)}}{l!} e^{-\frac{1}{2}\left(\frac{z_k - l - \eta_{w,k}}{\sigma_{w,k}}\right)^2} \right) \right)\tag{7}$$

where $z_1, \ldots, z_K$ denotes the single molecule image data, $\mu_{\theta,k} + b\Delta t$ denotes the mean pixel intensity, and $\sigma_{w,k}$ and $\eta_{w,k}$ denote the standard deviation and mean of the Gaussian noise component respectively [14].

In the absence of Gaussian noise, the maximum likelihood estimate was obtained by minimizing the negative of the function

$$\mathcal{L}(\theta|z_1,\ldots,z_K) := \sum_{k=1}^{K} \left((z_k) \cdot \ln(\mu_{\theta,k} + b\Delta t) - (\mu_{\theta,k} + b\Delta t) \right),$$ (8)

Function minimization was performed using the `fminunc` function provided in MATLAB's Optimization Toolbox.

In all calculations of standard deviations and means, only those estimates were used for which the MATLAB optimization functions successfully completed and the estimated location was within the image. The EstimationTool [23] was developed to facilitate the calculations involved in single molecule image fitting described by the equations above and supports both the maximum likelihood and nonlinear least squares algorithms.

Acknowledgements